

Article

Comparative Analysis of the Core Proteomes among the *Pseudomonas* Major Evolutionary Groups Reveals Species-Specific Adaptations for *Pseudomonas aeruginosa* and *Pseudomonas chlororaphis*

Marios Nikolaidis ¹, Dimitris Mossialos ², Stephen G. Oliver ³ and Grigorios D. Amoutzias ^{1,*}

¹ Bioinformatics Laboratory, Department of Biochemistry and Biotechnology, University of Thessaly, 41500 Larissa, Greece; marionik23@gmail.com

² Microbial Biotechnology-Molecular Bacteriology-Virology Laboratory, Department of Biochemistry and Biotechnology, University of Thessaly, 41500 Larissa, Greece; mosial@bio.uth.gr

³ Cambridge Systems Biology Centre & Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, UK; sgo24@cam.ac.uk

* Correspondence: amoutzias@bio.uth.gr; Tel.: +30-2410-565289; Fax: +30-2410-565290

Received: 22 June 2020; Accepted: 22 July 2020; Published: 24 July 2020

Abstract: The *Pseudomonas* genus includes many species living in diverse environments and hosts. It is important to understand which are the major evolutionary groups and what are the genomic/proteomic components they have in common or are unique. Towards this goal, we analyzed 494 complete *Pseudomonas* proteomes and identified 297 core-orthologues. The subsequent phylogenomic analysis revealed two well-defined species (*Pseudomonas aeruginosa* and *Pseudomonas chlororaphis*) and four wider phylogenetic groups (*Pseudomonas fluorescens*, *Pseudomonas stutzeri*, *Pseudomonas syringae*, *Pseudomonas putida*) with a sufficient number of proteomes. As expected, the genus-level core proteome was highly enriched for proteins involved in metabolism, translation, and transcription. In addition, between 39–70% of the core proteins in each group had a significant presence in each of all the other groups. Group-specific core proteins were also identified, with *P. aeruginosa* having the highest number of these and *P. fluorescens* having none. We identified several *P. aeruginosa*-specific core proteins (such as *CntL*, *CntM*, *PlcB*, *Acp1*, *MucE*, *SrfA*, *Tse1*, *Tsi2*, *Tse3*, and *EsrC*) that are known to play an important role in its pathogenicity. Finally, a holin family bacteriocin and a mitomycin-like biosynthetic protein were found to be core-specific for *P. chlororaphis* and we hypothesize that these proteins may confer a competitive advantage against other root-colonizers.

Keywords: *Pseudomonas*; core-proteome; phylogenomics; comparative genomics; species-specific adaptations

1. Introduction

The *Pseudomonas* genus is very diverse and ubiquitous and, thus far, it includes 272 species (<http://www.bacterio.net/pseudomonas.html>) isolated from many different environments, such as soil, water, air, and sediment, as well as from many types of hosts, such as animals, plants, fungi, and algae [1,2]. In addition, many of these species or strains are human, animal, or plant pathogens of great importance; in contrast, some actually promote plant growth. Of particular prominence is *Pseudomonas aeruginosa*, an opportunistic human pathogen that is resistant to many antibiotics. This species is a major cause of nosocomial infections and severely affects patients who are immunocompromised or suffering from cystic fibrosis [3]. Another close evolutionary relative, *Pseudomonas stutzeri*, is widely distributed in the environment and has been studied as a model for studies on denitrification, the degradation of pollutants, interactions with toxic metals, and, lately, as an opportunistic human pathogen especially in clinical settings [4]. *Pseudomonas syringae* is one of the most common and, accordingly, well-studied plant pathogens that infects the phyllosphere and lives on plant surfaces as an epiphyte [5]. Another prominent group within the genus is *Pseudomonas fluorescens*; this is no longer considered as an individual species but, rather, as a “species complex” or else a wider phylogenetic group that comprises at least 52 species [6,7]. Although various *P. fluorescens* strains are generally considered as soil and rhizosphere bacteria, some are also involved in human infections [7]. *Pseudomonas chlororaphis* is studied for its plant growth-promoting properties and especially for its production of secondary metabolites that protect the plant from fungi, nematodes, and insects [8]. *Pseudomonas putida* is a soil bacterium with a very diverse metabolism that is capable of degrading many organic compounds (including xenobiotics) and thus has been widely studied for use in bioremediation [9,10].

For many years, morphological and phenotypic characteristics were used to identify and differentiate the various species of this genus. Bacteria that were Gram-negative, strictly aerobic, and non-sporulating, and that appeared as motile bacilli were classified as *Pseudomonas* [11]. However, the new nucleic acid technologies of the late 20th century (especially the use of 16S ribosomal RNA (rRNA) sequences, by Carl Woese) permitted a more refined and deeper classification within the gamma-Proteobacteria [12]. Such 16S rRNA studies helped to identify species that were mis-assigned to the *Pseudomonas* genus when classification was based solely on phenotypic/biochemical traits [11]. In addition, 16S rRNA studies in parallel with more modern methodologies have allowed the re-classification of some bacteria as members of this genus [13].

Although 16S rRNA sequencing is well-established and still a widely approved/used method for classifying a new species within this genus, it does not provide enough resolution for accurate classification at the species level [14,15]. Thus, the MultiLocus Sequence Typing (MLST) approach based on several protein-coding genes has been implemented to define new *Pseudomonas* species, with the help of the PseudoMLSA database [16]. However, the MLST approach has its limitations, since strains of the same sequence type (MLST) have been observed to be genetically distant at the whole-genome level [17]. Recently, the authors of [2] provided two phylogenetic trees of the various *Pseudomonas* species, one based on the 16S rRNA and the second based on multilocus sequence analysis (MLSA) of the concatenated 16S rRNA, *gyrB*, *rpoB*, and *rpoD* genes. In the later tree, they indicated the groups and phylogenetic lineages that have been proposed by earlier studies [6,18]. Furthermore, the most recent taxonomic analysis of 217 species type strains of the genus was based on a four-gene analysis [19]. An even more advanced version of the MLST approach, named rMLST, is based on 53 ribosomal proteins that are present in most bacteria, are distributed around the genome, and are under stabilizing selection [20].

However, the ever-decreasing cost of sequencing an entire bacterial genome has given rise to the new era of phylogenomics, where many diverse bacterial strains are sequenced and their taxonomy is determined on the basis of the phylogenies of hundreds or even thousands of orthologous genes/proteins present in all of them [21–23]. Although horizontal gene transfer (HGT) is a major evolutionary force in prokaryotes that could potentially confound such studies [24], the majority of protein families are transmitted by vertical inheritance [25,26]. A recently revised bacterial taxonomy of 94,759 genomes was based on 120 ubiquitous single-copy proteins [27]. Thus, the noise due to phylogenetic incongruities that may arise due to limited recombination events or HGT becomes negligible against the real evolutionary signal of the remaining hundreds or thousands of orthologous groups. These phylogenomic approaches rely on the concept of the pangenome of a defined set of organisms that consists of core, dispensable, and strain-specific genes [17,28]. In addition, it has been proposed to ease the criterion for core genes/proteins that are present in 95% of the analyzed strains, so as to account for sequencing errors [29]. For pangenomic analyses, it has been suggested that at least five genomes are needed [17].

The *Pseudomonas* Genome Database is a great resource that organizes data from more than 9000 *Pseudomonas* genomes at various levels of genome assembly [30]. In addition, a large-scale sequencing effort has focused on the widely studied human pathogen *P. aeruginosa* [31,32]. With all these genomes available, new computational approaches have been introduced in order to define new *Pseudomonas* species by using the average nucleotide identity (ANI) at a threshold of 94–96% ANI similarity or the genome-to-genome distance (GGDC) [33–35].

The goal of the present study was to analyze the available high-quality complete *Pseudomonas* genomes/proteomes via a phylogenomic approach in order to better understand the extent to which the various species of the genus; the major evolutionary groups that they form; and the properties of the core genome/proteome at the genus, group, and species levels are well defined.

2. Materials and Methods

A total of 8999 proteomes (accessed on October 2019) were downloaded from the *Pseudomonas* Genome Database. One *Candidatus* proteome was subsequently excluded. In addition, if two or more assemblies existed for the same strain, only the one with the highest sequence coverage was kept. Thus, we retained 8883 proteomes. Of these, 494 were annotated as “complete”, whereas, for the rest, their assembly level was characterized as either “chromosome”, “scaffold”, or “contig”. In order to ensure that our analyses were not affected by badly assembled genomes/proteomes that could result in a significant underestimation of the core *Pseudomonas* proteome, we focused on the 494 “complete” proteomes.

In order to identify the core proteome of a set of organisms, we developed three Python scripts that perform best reciprocal protein Basic Local Alignment Search Tool (BLAST), using the biopython blast+ tools. The user has to define a reference proteome that is used to perform best reciprocal BLAST against all the other proteomes of the set. The parameters of the best reciprocal BLAST are $E\text{-value} < 1 \times 10^{-5}$. The user may define another $E\text{-value}$ threshold. We also implemented an additional criterion in order to stringently filter out problematic orthologues. For all reciprocal BLAST hits of the reference proteome against a selected proteome, the Python script gathers all the percent identities (of the BLAST results), estimates the mean value and standard deviation, and then filters out all best reciprocal BLAST hits that have identities two standard deviations below the average value. The user may alter this threshold of standard deviations. This approach defines a cut-off of orthology that is adjustable, depending on the

genetic distance of the two genomes/proteomes undergoing reciprocal BLAST, instead of fixed cut-offs of sequence identity/similarity or defined BLAST score ratios, as applied in many previous pangenome analyses [28,36–39]. We did not implement any additional criterion/threshold of a required minimum pairwise alignment length. Afterwards, a matrix of best reciprocal BLAST hits is generated, where the rows correspond to the proteins of the reference proteome and their orthologues in the other proteomes, and the columns correspond to the various proteomes of the selected set. Next, a second matrix is generated that includes only the proteins and the orthologues that are present in all of the analyzed proteomes. This second matrix corresponds to the actual core proteome. Afterwards, the script generates multiple alignments with the Muscle software [40] for all identified groups of core orthologues, concatenating them in a super-alignment and then filtering badly aligned regions with the gblocks software (default parameters). The filtered alignment is then used to generate a neighbor-joining phylogenomic tree with the Kimura model and 500 bootstraps [41]. All the above parameters and directories that contain the reference proteome and the other proteomes are defined in a parameters file that is used by the scripts. The scripts are available upon request.

In order to identify the core proteome of the whole genus, we used (as a reference) the model strain *P. aeruginosa* PAO1. The same strain was used as a reference for identifying the core proteome of the *P. aeruginosa* species. However, in order to identify the core proteomes of the other evolutionary groups, we used a specific reference strain from within each group. Our approach of using a reference proteome for identifying core proteins with reciprocal BLASTs is scalable and thus may be readily applied in the future. The number of reciprocal BLASTs grows linearly with the total number of genomes/proteomes analyzed. In contrast, an all vs. all reciprocal BLAST approach is subject to a combinatorial explosion and so is not scalable. Future analyses with the all vs. all approach may be expected to require vast amounts of computing power, given the rate at which bacterial genomes are now being sequenced.

Average nucleotide identity for species boundary identification [34], based on the MUMmer software (from now on abbreviated as ANIm), was calculated with the pyani software [42]. In order to define the borders of a certain *Pseudomonas* species or group, (i.e., *P. aeruginosa*, *P. stutzeri*, *P. putida*, *P. fluorescens*, *P. chlororaphis*), we observed the genus phylogenomic tree and the ANIm values of the various genomes against a selected strain of a particular species. A workflow is presented in Figure 1.

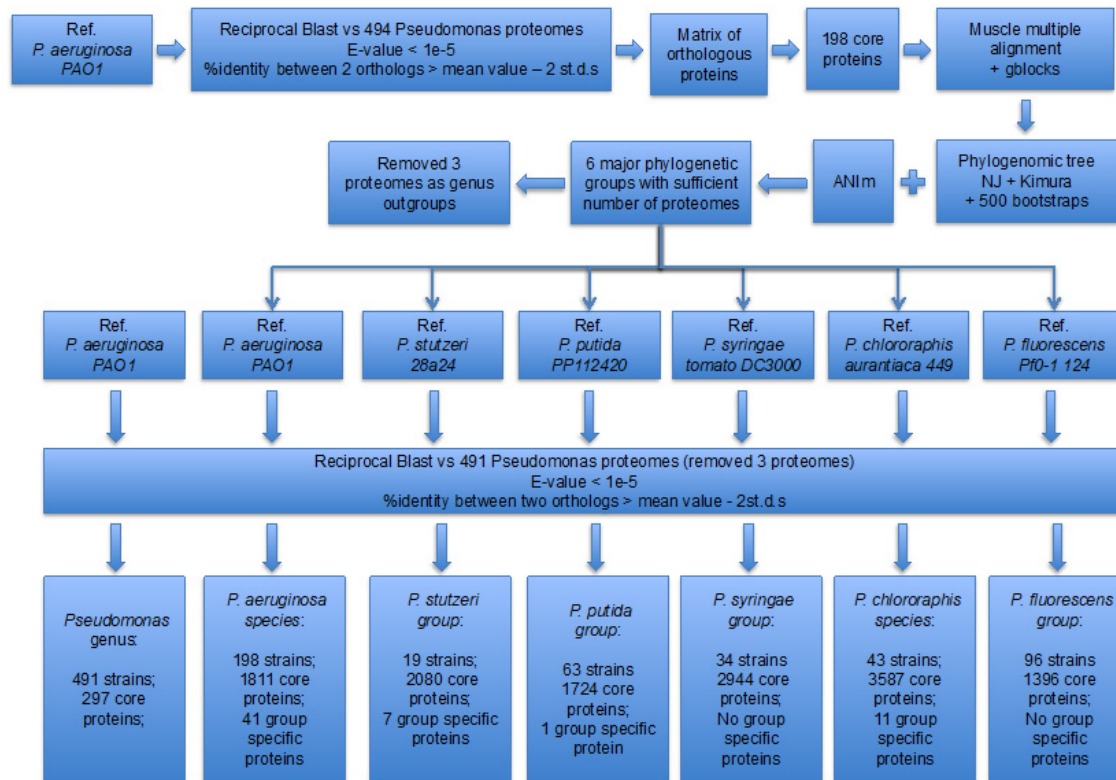


Figure 1. Workflow of the phylogenomic and core proteome analyses.

In order to compare the results from our approach with those from other published computational tools, we also implemented Orthofinder [43], which follows an all vs. all pairwise alignment approach, such as BLAST, or other fast alternative methods, such as Diamond [44]. For Orthofinder, we provided the rooted species tree that we calculated on the basis of our approach and we applied all other default parameters. Orthofinder analyzed the 494 proteomes in 3 days in an Intel dual processor workstation with 36 threads. Although it is feasible to run such all vs. all analyses with 500–1000 proteomes, due to the exponential computational cost of an all vs. all approach, such an analysis would be infeasible if we repeated it for the 8999 originally downloaded proteomes (estimated run time is 1026 days).

The assignment of functional category of the identified core proteins was based on the EGGNOG database v4.5 [45] and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthologies with the help of the KAAS tool [46], as well as on COG [47]. Heatmaps were generated with the help of the R programming language (<https://www.r-project.org/>).

3. Results and Discussion

3.1. Identification of the Major Phylogenetic Groups and Species

We analyzed all the 494 complete *Pseudomonas* proteomes using as reference the proteome of the first completely sequenced strain, that of *P. aeruginosa* PAO1 [48]. In this set of complete proteomes, the majority of them were annotated as *P. aeruginosa* strains (187), whereas the other most numerous proteomes were strains annotated as *P. chlororaphis* (43), *P. syringae* (28), *P. putida*

(27), *P. fluorescens* (25), and *P. stutzeri* (18). Our first analysis identified 198 core proteins for the whole genus (see Supplementary File 1, Table S1). The multiple alignment of these 198 core proteins contained 27,997 variable sites (after G-blocks filtering) that were used to generate a phylogenomic tree in Seaview4 [41] with neighbor-joining (Kimura model) and 500 bootstraps (see Figure 2 and Figure S1).

However, we observed an oddly long branch for *Pseudomonas mesoacidophila* ATCC 31433 7131, which is in actuality a mis-assigned *Burkholderia cepacia* [49]. Thus, this particular strain was further used as outgroup. In addition, at the base of this phylogenomic tree, we also observed two branches/strains that raised doubts about their correct classification as members of the genus. Accordingly, on the basis of the phylogenomic tree, we removed several basal branches/strains in a step-wise manner, created five genus core sets (see Figure 2 and Supplementary File 1; Table S1), and observed how the number of core proteins increased. Our rationale was that removal of strains that are not members of the genus would significantly increase the number of core proteins, whereas once we reached the genus borders, removal of true genus strains would not cause a significant increase in the number of core proteins. Indeed, we observed a significant increase of core proteins until core set 4, which contained 297 core proteins and excluded 3 *Pseudomonas* annotated strains, *P. mesoacidophila* (the mis-annotated *Burkholderia*), *Pseudomonas* sp. S-6-2 (PubMLST assigned it as *Pseudomonas*), and *Pseudomonas* sp. QZS01 (PubMLST did not assign it as *Pseudomonas*). In addition, *Pseudomonas* sp. QZS01 had an atypically low GC content of 38% with only 3004 proteins, whereas the GC content for all the other strains ranged between 54 and 67%. However, core set 5 further excluded four more strains—two *Pseudomonas psychrotolerans*, one *Pseudomonas oryzihabitans*, and *Pseudomonas* sp. LTJR-52—and contained 305 proteins, an increase of only eight proteins from core set 4. In addition, PubMLST annotated these four strains as *Pseudomonas* (*Pseudomonas* sp. LTJR-52 was annotated by PubMLST as *Pseudomonas luteola*). Thus, we subsequently used core set 4 as the genus core set.

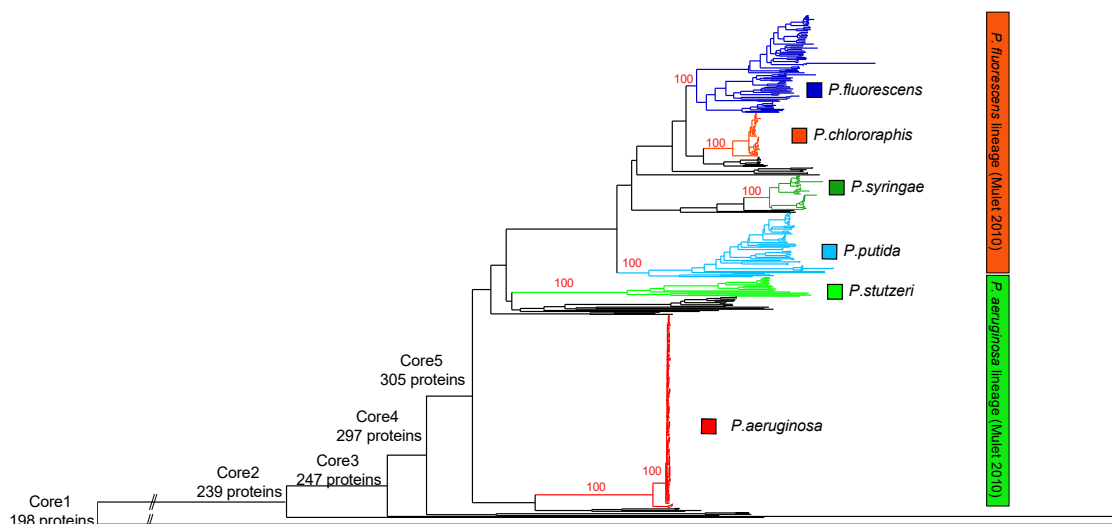


Figure 2. The phylogenomic neighbor-joining tree of the 494 complete proteomes. The tree was based on 198 core proteins (core set 1), using the Kimura model and 500 bootstrap values. The various genus core sets are indicated on the tree.

According to a Gene Ontology enrichment analysis with PANTHER [50], the core proteins were housekeeping proteins, mostly related to metabolism, organic compound biosynthesis,

translation, and gene expression, as expected (see supplementary file 1, Table S2), in accordance with the work of [51].

The observation of the phylogenomic tree together with ANIm values from selected representatives of key species and the PubMLST database allowed us to identify two well-defined species, *P. aeruginosa* and *P. chlororaphis*, as well as four other groups, *P. fluorescens*, *P. putida*, *P. stutzeri*, and *P. syringae*, but with much higher intragroup diversity. Our phylogenomic analysis is in accordance with a previous analysis of 107 *Pseudomonas* strains, using the 16S rRNA, *gyrB*, *rpoB*, and *rpoD* [6]. This previous analysis identified two major lineages or intrageneric groups (IGs). The first IG contained the species *P. aeruginosa*, *P. stutzeri*, and *Pseudomonas oleovorans*. However, the second IG was more complex and contained six groups, represented by *P. fluorescens*, *P. syringae*, *Pseudomonas lutea*, *P. putida*, *Pseudomonas anguilliseptica*, and *P. straminea*. The *P. fluorescens* group was further divided into nine subgroups, one of which was *P. chlororaphis*. They also included *Pseudomonas psychrotolerance*, *P. oryzihabitans*, and *P. luteola* within the genus, but not in any IG. The genus-level phylogenomic tree is important for identifying mis-assigned *Pseudomonas* genomes, a problem that has been observed in the past [35,52]. Our phylogenomic tree also revealed the position of *Pseudomonas* sp. S-6-2 and *Pseudomonas* sp. QZS01 at the borders of the genus. A recent study identified 13 *Pseudomonas* groups, composed of 2 to 66 species each [51].

As is evident from Figure 2 and Figure S1, the *P. aeruginosa* species (composed of 189 strains) is well grouped and monophyletic, in accordance with previous observations [6,29]. One genome was misidentified as *P. fluorescens* NCTC10783 9425 and should be renamed as a *P. aeruginosa* strain. This is also supported by the ANIm values when we used *P. aeruginosa* PAO1 as the reference strain, as well as by PubMLST. Interestingly, four *P. aeruginosa* strains (PA7 119, CR1 6461, AR441 6528, AR 0356 7083) that seem to form the most distant subgroup have ANIm values around 94%, whereas all the other members of the group have ANIm values above 98%. However, the inclusion of these four strains into the *P. aeruginosa* species is well supported, not only by the phylogenetic tree, but also from the PubMLST analysis. The core proteome of the *P. aeruginosa* species comprised 1811 proteins. A separate phylogenomic analysis of these 189 proteomes (see Figure S1), based on the 1811 *P. aeruginosa* core proteins and 31,145 variable sites, further confirmed the mis-annotation of *P. fluorescens* NCTC10783 9425. The protein count and GC content of the strains of this group range between 5500–7352 (average: 6192) and between 65.6–66.9% (average: 66.1%), respectively. Although it is a very well-defined species, phylogenomically and in terms of ANIm values, it is surprisingly diverse in terms of protein content, thus revealing a very dynamic accessory proteome, in accordance with several analyses [53–55]. In particular, an analysis of 1311 *P. aeruginosa* genomes revealed 665 core genes, a five-group population structure, and evidence of very extensive horizontal gene transfer [55], whereas another analysis of 22 strains showed that up to 20% of the genome consists of transferable genetic loci [54]. Another analysis of 17 reference strains revealed a core set of 5233 orthologs [56]. In addition, it appears that, on average, industrial strains have the largest genomes, followed by environmental strains, and then clinical isolates [57]. Interestingly, a recent study based on experimentally determined essential genes (by using transposon insertion sequencing) for several *P. aeruginosa* clinical isolates across different conditions revealed 321 core essential genes [58], with 24% (76/321) and 58% (185/321) of them being members of the genus and *P. aeruginosa* core sets of our analysis (see supplementary file 1, Table S10). In total, five studies of this type have identified 864 essential genes [58–62], with 15% (128/864) and 42% (359/864) of them being members of the genus and *P. aeruginosa* core sets of our analysis.

Another well-defined *Pseudomonas* species, based on the genus phylogenomic tree and the ANIm values was *P. chlororaphis*, comprised of 43 genomes. *Pseudomonas* sp. MRSN12121 3055 was found to be within the *P. chlororaphis* species, whereas *P. chlororaphis* UFB2 isolate Soil 2921 was not within the *P. chlororaphis* group that we defined but was within our wider *P. fluorescens* group. The core proteome of the *P. chlororaphis* species, using *P. chlororaphis* subsp. *aurantiaca* 449 7539 strain as the reference point, comprised 3587 proteins. A separate phylogenomic analysis (see Figure S1) of these 43 proteomes, on the basis of their 3587 core proteins and 103483 variable sites, further confirmed the inclusion of *Pseudomonas* sp. MRSN12121 3055 within the *P. chlororaphis* species. The protein count and GC content of the strains of this group range between 5599–6401 (average: 6076) and between 61.9–64% (average: 62.8%), respectively.

The *P. fluorescens* group that we defined, on the basis of the genus phylogenomic tree, comprised 96 genomes and displayed high levels of phylogenetic heterogeneity, in accordance with [29]. It comprised many species, such as *Pseudomonas corrugata*, *Pseudomonas brassicacearum*, *Pseudomonas frederiksbergensis*, *Pseudomonas mandelii*, *Pseudomonas kribbensis*, *Pseudomonas koreensis*, *Pseudomonas mucidolens*, *Pseudomonas veronii*, *Pseudomonas antarctica*, *Pseudomonas azotoformans*, *Pseudomonas trivialis*, *Pseudomonas lurida*, *Pseudomonas azotoformans*, *Pseudomonas poae*, *Pseudomonas libanensis*, *Pseudomonas synxantha*, and *Pseudomonas orientalis*. Interestingly, the genomes that were named as *P. fluorescens* did not form a monophyletic clade within this wider group, but were dispersed. In addition, two genomes within the *fluorescens* group have been mistakenly named as *P. stutzeri* (KGS-2 9480, KGS-8 9481), one genome was mistakenly named as *P. putida* (JBC17 6779), one was mistakenly named as *P. syringae* (isolate inb918 11494), and another as *P. chlororaphis* (UFB2 isolate Soil 2921). In essence, the name *P. fluorescens* encompasses many different and diverse species lumped together. The core proteome of the *P. fluorescens* group, using *P. fluorescens* Pf0-1 124 strain as the reference point, comprised 1396 proteins. A separate phylogenomic analysis (see Figure S1) of these 96 proteomes, on the basis of their 1396 core proteins and 115,099 variable sites further confirmed the observations above. The protein count and GC content of the strains of this group range between 4152–6678 (average: 5603) and between 58.7–62% (average: 60.3%), respectively. Other analyses, based on 3 and 8 *P. fluorescens* genomes, revealed 3642 [63] and 3039 [64] core genes, respectively. In addition, a comparative genomic analysis of 71 *P. fluorescens* genomes identified eight major subgroups and developed a set of nine genes as markers for classification within this lineage [65].

The *P. syringae* group that we defined, on the basis of the genus phylogenomic tree, comprised 34 genomes. It also included species such as *Pseudomonas avellanae*, *Pseudomonas savastanoi*, *Pseudomonas amygdali*, and *Pseudomonas cerasi*. The genomes that were named as *P. syringae* did not form a monophyletic clade within this wider group, but were dispersed. In addition, the *P. syringae* group formed three subgroups, according to our phylogenomic analysis and the ANIm values. The core proteome of the *P. syringae* group, using *P. syringae* pv tomato DC3000 111 strain as the reference point, comprised 2944 proteins. A separate phylogenomic analysis of these 34 proteomes (see Figure S1), based on their 2944 core proteins and 110,643 variable sites, further confirmed the above observations. The protein count and GC content of the strains of this group range between 4973–6026 (average: 5465) and between 58–59.3% (average: 58.6%), respectively.

The *P. putida* group that we defined, on the basis of the genus phylogenomic tree, comprised 63 genomes. It also included species such as *Pseudomonas alkylphenolia*, *Pseudomonas monteilii*, *Pseudomonas cremoricolorata*, *Pseudomonas fulva*, *Pseudomonas parafulva*, *Pseudomonas entomophila*, *Pseudomonas mosselii*, and *Pseudomonas plecoglossicida*. The genomes that were named as *P. putida* did not form a monophyletic clade within this wider group, but were dispersed. The core

proteome of the *P. putida* group, using *P. putida* PP112420 5661 strain as the reference point, comprised 1724 proteins. A separate phylogenomic analysis of these 63 proteomes (see Figure S1), based on their 1724 core proteins and 155,470 variable sites, further confirmed the above observations. The protein count and GC content of the strains of this group range between 3748–6780 (average: 5197) and between 58.7–64.4% (average: 62.3%), respectively. Another analysis that was based on nine strains identified 3386 core genes [66].

The *P. stutzeri* group that we defined, on the basis of the genus phylogenomic tree, comprised 19 genomes. It also included *Pseudomonas balearica*, *Pseudomonas* sp. 20 BN 3032, and *Pseudomonas* sp. R2A2 5971. Although the *P. stutzeri* group is well defined, the ANIm values and the branch lengths of the genus phylogenomic tree support the idea that this is not a species, but rather a wider taxonomic group. The core proteome of the *P. stutzeri* group, using *P. stutzeri* 28a24 3065 strain as the reference point, comprised 2080 proteins. A separate phylogenomic analysis of these 19 proteomes (see Figure S1), based on their 2080 core proteins and 193,427 variable sites, further confirmed the above observations. The protein count and GC content of the strains of this group range between 3342–4524 (average: 4086) and between 60.3–64.7% (average: 63.2%), respectively.

3.2. Properties of the Core Proteomes

A critical issue that affects the estimation of a core genome/proteome is the number of strains used to analyze it. As more proteomes are included in the analysis, less proteins remain present in all the analyzed strains. By observing how the curve of core proteins decreases as more strains are analyzed, one can estimate the true number of core proteins for that set of organisms. We generated a graph by random sampling (20 times each) 16, 30, 50, 100, 150, up to 491 *Pseudomonas* strains and estimated the average number of core proteins for that sampling depth of the genus. Curve fitting by applying exponential decay suggests that our sampling depth provides a set of proteins that is very close to the core genus proteome (see Figure 3). In accordance with these observations, our genus core set (based on 491 proteomes—core set 4) composed of 297 proteins. Previous analyses based on 14, 166, and 1073 diverse *Pseudomonas* genomes had identified 1705 core genes [63], 794 core genes [51], and 1224 protein-coding gene families [29], respectively. Another type of analysis that was based on 58 complete and 432 draft *Pseudomonas* genomes identified a core set consisting of 2687 and 726 protein domains [67].

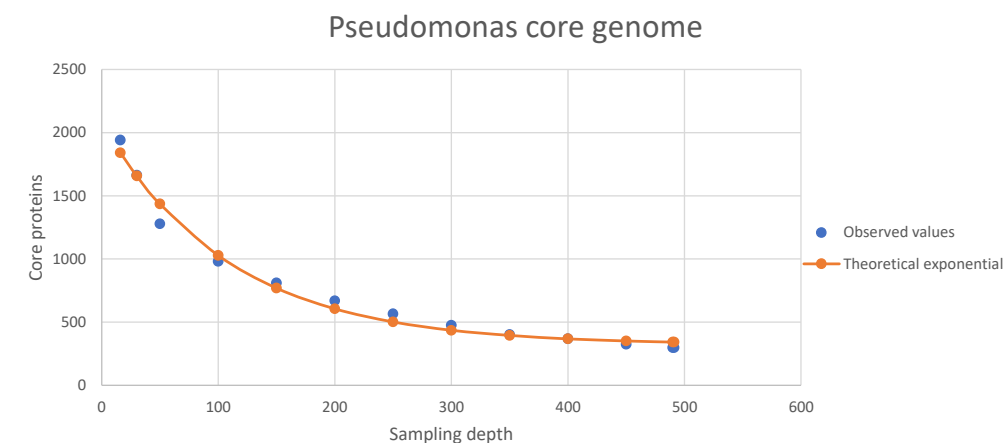


Figure 3. The *Pseudomonas* genus core proteome, depending on the sampling depth (i.e., number of proteomes analyzed).

We repeated this analysis for the *P. aeruginosa*, *P. chlororaphis*, *P. fluorescens*, *P. syringae*, and *P. stutzeri* groups. The core proteome of *P. aeruginosa*, with 189 analyzed proteomes, seems to approach saturation, whereas this is not the case for the other groups. Most probably, the saturation level of a certain core proteome depends on the sampling depth (i.e., the number of sequenced genomes analyzed), the quality of the data (i.e., quality of the genome assembly), and the level of diversity within that particular group.

For each of these seven core proteome sets, that is, of the whole genus and the six evolutionary groups, we identified the frequency of functional categories of their core proteins using the EggNog 4.5 server. The results are summarized in Table 1 and in Supplementary File 1, Table S9. Accordingly, Gene Ontology analysis of the genus core proteome revealed a high enrichment for metabolism (e.g., nitrogen compound metabolic process) translation and gene expression, as expected (see Supplementary File 1, Table S2). It is intriguing that many proteins of the various core sets are of unknown function, thus highlighting the need for more focused in-depth molecular biology/biochemistry analyses by wet-lab biologists.

Table 1. Summary table of the core proteomes and their functional categories for the *Pseudomonas* genus and the six evolutionary groups. Note that, for the phylogenomic analysis of the entire genus (core4), we used the alignment produced by core1, that also includes 3 outgroups.

	Genus Core4	<i>Pseudomonas</i> <i>aeruginosa</i>	<i>Pseudomonas</i> <i>chlororaphis</i>	<i>Pseudomonas</i> <i>stutzeri</i>	<i>Pseudomonas</i> <i>putida</i>	<i>Pseudomonas</i> <i>fluorescens</i>	<i>Pseudomonas</i> <i>syringae</i>
Number_of_Strains	491	189	43	19	63	96	34
Amino acids in phylogenomic alignment	Core1:27,997	31,145	103,483	193,427	155,470	115,099	110,643
Core proteins	297	1811	3587	2080	1724	1396	2944
% core proteins with significant presence in other groups (≥90% presence)	100.0	44.3	38.7	70.1	70.5	62.7	50.1
Group-specific core proteins	-	41	11	7	1	0	0
Relaxed group-specific core proteins (10% in others)	-	84	32	32	4	0	61
Relaxed group-specific core proteins (20% in others)	-	116	61	51	4	0	87
%core—Unknown	19.9	33.4	30.4	27.1	25.4	26.6	31.3
%core—Other	24.9	16.6	20.1	19.9	17.9	17.1	18.8
%core—K:Transcription	4.7	7.3	7.8	4.6	5.4	6.5	5.9
%core—E:Amino acid transport and metabolism	11.4	7.2	8.6	7.2	9.5	10.7	8.2
%core—P:Inorganic ion transport and metabolism	4.4	5.9	5.5	4.5	5.5	4.7	5.8
%core—C:Energy production and conversion	4.4	6.5	5.8	6.3	7.0	6.1	4.8
%core—M:Cell wall/membrane/envelope biogenesis	6.7	4.4	5.5	6.1	5.3	5.2	5.7
%core—J:Translation, ribosomal structure and biogenesis	13.1	5.6	4.2	6.9	7.1	6.0	5.2
%core—H:Coenzyme transport and metabolism	7	3.4	3.5	4.6	5.0	5.1	4.2
%core—L:Replication, recombination and repair	3.4	3.0	2.8	5.1	4.4	4.0	3.7

We investigated the overlap of orthologues among the various core proteomes of the genus and the six evolutionary groups by mapping them to KEGG orthologs with KAAS. We also identified their wider functional category with EggNOG 4.5. The vast majority of the genus core proteins were mapped to KEGG orthologues, whereas for each of the other six core proteomes, between 50–70% of their proteins were mapped. Thus, more than 2000 KEGG orthologous groups were identified that had an orthologue in one or more core proteome sets. Of note, this analysis should be interpreted with caution, because the core proteomes of *P. chlororaphis*, *P. fluorescens*, *P. putida*, *P. stutzeri*, and *P. syringae* are based on a small number of complete proteomes and show no evidence of saturation, in contrast to *P. aeruginosa*. In addition, every phylogenetic group has a different level of intragroup diversity. Overall, the vast majority (91–99.5%) of the core proteins of each group were also identified in one or more of the other core sets.

We further investigated to what extent the core proteins of a certain group had a very significant presence in all the other evolutionary groups (see Table 1). Our criterion was that the core protein of that set should be present at least in 90% of the strains in each of the other five evolutionary groups. We observed such a high presence for 44% (803/1811) of the core proteins in *P. aeruginosa*, 39% (1387/3587) of the core proteins in *P. chlororaphis*, 70% (1458/2080) of the core proteins in *P. stutzeri*, 70% (1215/1724) of the core proteins in *P. putida*, 63% (875/1396) of the core proteins in *P. fluorescens*, and 50% (1475/2944) of the core proteins in *P. syringae*. By obtaining the functional annotation of these widely distributed core proteins via eggNOG, we observed that they were enriched (more than twofold compared to the background frequency of the *P. aeruginosa* PAO1 genome) for functions related to translation and the ribosome, as well as co-enzyme and nucleotide transport and metabolism.

Next, we identified 60 group-specific core proteins that may play a critical role in the adaptation of that specific group. Our criterion was that these proteins were found in all strains of that specific group and were absent in all the other *Pseudomonas* strains (based on our strictly defined orthology criteria of best reciprocal BLAST hit). These group-specific core proteins were usually hypothetical or of unknown function. More specifically, we observed 41 out of the 1811 core orthologues of the *P. aeruginosa* species that were specific to that group (see Supplementary File 1, Table S3), with 26 of them being annotated as hypothetical. Accordingly, the same analysis revealed 11 out of the 3587 core orthologs of the *P. chlororaphis* species that were specific for that group (see Supplementary File 1, Table S4). For the *P. stutzeri* group, we observed that 7 out of its 2080 core proteins were group-specific (see Supplementary File 1, Table S5). For *P. putida*, only 1 out of the 1724 core proteins was group-specific (see Supplementary File 1, Table S6). For the *P. fluorescens* and *P. syringae* groups, none of their 1396 and 2944 core orthologs (respectively) were group-specific (see Supplementary File 1, Tables S7 and S8).

In order to assess the reliability of our methodology, we also identified group-specific core proteins with Orthofinder (see Supplementary File 1, Table S11). Orthofinder correctly identified 41 of our 60 group-specific core proteins. It did not identify 14 of our group-specific core proteins because the Orthofinder orthogroups were broader and also included within them other widely distributed paralogues (of the reference proteome) and their respective orthologues. Orthofinder also did not identify five of our group-specific core proteins, because within those orthogroups, Orthofinder further included distant homologues from out-of-the-group species that did not pass our orthology criteria. Finally, Orthofinder identified three group-specific orthogroups that we did not detect. In these cases, our orthology groups further included out-of-the-group orthologues. However, Orthofinder would assign these out-of-the-group orthologues in other orthogroups. Thus, we conclude that our approach is accurate and scalable for future analyses with thousands of proteomes.

By further investigating Supplementary File 1, Tables S3–S8, one can identify more group-specific core ortholog candidates after applying two more relaxed criteria, where the group-specific core protein is allowed to be present in other groups as well, but with a maximum presence of 10% or 20% (in each group). By applying these two relaxed criteria of 10% and 20%, the group-specific core proteins are, respectively, 84 and 116 in *P. aeruginosa*, 32 and 61 in *P. chlororaphis*, 32 and 51 in *P. stutzeri*, 4 and 4 in *P. putida*, 0 and 0 in *P. fluorescens*, and 61 and 87 in *P. syringae*. The lack of group-specific core proteins for the *P. fluorescens* group could be attributed to the fact that this is a very diverse group. In addition, one may further hypothesize that this great diversity combined with the lack of group-specific core proteins could be attributed to the fact that this may be the oldest evolutionary group that gave rise to the other more specialized evolutionary groups; however, its position in the genus phylogenomic tree does not support this hypothesis. The relatively high number of group-specific core proteins in the *P. aeruginosa* group could be attributed to the fact that this is a rather specialized species that frequently uses humans as hosts and has very low genetic diversity (albeit having a very dynamic accessory proteome), whereas the other groups are not specialized as mainly human pathogens. An alternative explanation is that there is a sampling bias in favor of human-pathogenic *P. aeruginosa* strains.

3.3. Many *P. aeruginosa*-Specific Core Proteins Contribute to Pathogenicity

Our comparative analysis identified several *P. aeruginosa*-specific core proteins (present in all *P. aeruginosa* strains, but not present at all in any other *Pseudomonas* strains) with an important role in its pathogenicity. Our findings, explained further below, are summarized in Table 2. Notably, two of these (*cntL* and *cntM*) are members of a four-gene operon (*cntOLMI*), which is regulated by zinc levels. This operon is responsible for the production of pseudopaline, an opine metallophore used to scavenge metals such as nickel and zinc from the host that is believed to be implicated in pathogenesis [68,69]. Homologous operons that produce such opine metallophores are associated with virulence and have also been identified in other human pathogens, such as *Staphylococcus aureus*, *Serratia marcescens* (WW4), and *Yersinia pestis* (CO92). It has been shown that opine metallophores of this kind are probably implicated in nickel, cobalt, and iron acquisition in metal-limited environments [69,70]. Interestingly, cobalt is essential for the formation of biofilms, under oxygen-limited conditions, by *P. aeruginosa* [71]. The other two members of this operon encode membrane proteins implicated in pseudopaline export (*cntI*:PA4834) and the import of pseudopaline-metal complexes (*cntO*:PA4837) [68]. These two coding sequences were absent in only two and four *P. aeruginosa* strains, respectively. Apart from pseudopaline, pyoverdine, and pyochelin, there are the other two major metallophores (siderophores) identified in *P. aeruginosa*. Pyoverdine consists of three distinct structural parts: a conserved chromophore, a variable peptide arm, and a small dicarboxylic acid connected via an amide group to the chromophore. Many genes are required for pyoverdine biosynthesis, but four genes (*pvdL*, *pvdI*, *pvdJ*, *pvdD*) encoding non-ribosomal peptide synthetases (NRPSs) are the major biosynthetic genes [72,73]. NRPSs are large multimodular enzymes that enable microorganisms to produce a large repertoire of secondary metabolites including antibiotics, toxins, immunomodulators, and siderophores [74,75]. Pyoverdine biosynthesis starts with PvdL, which incorporates L-Glu, D-Tyr, and L-Dab (2,4-diaminobutyrate) into a precursor peptide. Through condensation and further modifications, this peptide forms the chromophore [76]. The *pvdL* gene was identified in 90% (171/189) of *P. aeruginosa* strains and in 84% (255/305) of all the other *Pseudomonas* strains in this study. Pyochelin biosynthesis requires salicylic acid and two cysteine residues. Three genes (*pchD*, *pchE*, *pchF*) are involved in pyochelin synthesis, the last two (*pchE*, *pchF*) encoding NRPSs [72]. The pyochelin synthetase PchF (PA4225) was present in 93% (176/189) of *P. aeruginosa* strains

and in 19% (58/305) of all the other *Pseudomonas* strains. Both pyoverdine and pyochelin contribute to the pathogenicity exerted by *P. aeruginosa* in lungs of patients with cystic fibrosis [72]. Our analyses show that, overall, pseudopaline is *P. aeruginosa*-specific, phyochelin is mostly limited to *P. aeruginosa*, whereas pyoverdine has a very wide distribution across the genus.

PlcB (PA0026) and Acp1 (PA1869) are another two *P. aeruginosa*-specific core-proteins, with an important role in motility and pathogenesis. More specifically, PlcB is a phospholipase C implicated in directional twitching motility during chemotaxis [77]. PlcB expression is under the control of quorum sensing, a strategy employed by many pathogens during infection [78]. Regarding the potential role of PlcB in pathogenesis, it may be noted that the levels of phosphatidylethanolamine (the substrate of the PlcB protein) are increased in the bronchoalveolar lavage fluid of young adults with cystic fibrosis (CF) in comparison with age-matched controls without CF [79]. Consequently, it is plausible that these increased levels of phospholipids could serve as a chemoattractant or a factor enhancing chemokinesis for *P. aeruginosa* bacteria that have initially colonized the upper airways in CF patients [77]. Acp1 is the only one of three acyl carrier proteins that functions in fatty acid synthesis. Replacement of *Acp1* caused reduced production of the quorum-sensing signals (N-acyl homoserine lactones) and abolished swarming motility, thus demonstrating its indirect role in the pathogenicity exerted by *P. aeruginosa* [80].

Our analyses further identified two interesting *P. aeruginosa*-specific core-proteins, MucE (PA4033) and SrfA (PA2559), that are involved in mucilage production (mucoidy). *P. aeruginosa* very often exhibits mucoidy, a phenotype that is due to the overproduction of an exopolysaccharide called alginate. The emergence of mucoid *P. aeruginosa* isolates in sputum specimens from CF patients signifies the onset of chronic respiratory infections. Mucoidy plays an important role in the pathogenesis of *P. aeruginosa* infections in CF patients, which might include increased resistance to antibiotics, increased resistance to phagocytic killing, and assistance in evading the host's immune response. MucE is a small envelope lipoprotein that positively regulates mucoid conversion, thus directly contributing to pathogenicity [81,82].

Table 2. Summary of important group-specific and related proteins in *P. aeruginosa* and *P. chlororaphis*.

Orthologue					Closest Parologue in Reference Strain				Closest Homologue outside of Group			
Organism	Name	General Function	Presence within Group	Presence outside of Group	Accession	Name	Annotation	% Identity	Organism	Name	Annotation	% Identity
<i>P. aeruginosa</i>	cntO	pseudopaline production	185/189	234/305	PA0151	-	probable TonB-dependent receptor	34.84	<i>Pseudomonas furukawaii</i>	KF707_C_RS00865	TonB-dependent receptor	34.56
<i>P. aeruginosa</i>	cntL	pseudopaline production	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	cntM	pseudopaline production	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	cntI	pseudopaline production	187/189	2/305	PA2628	-	hypothetical protein	34.17	<i>Pseudomonas mendocina</i>	EL191_RS11685	DMT family transporter	34.52
<i>P. aeruginosa</i>	pvdL	pyoverdine biosynthesis	171/189	255/305	PA2400	pvdJ	pvdJ	47.28	<i>Pseudomonas fulva</i>	CJ462_RS07705	amino acid adenylation domain-containing protein	39.98
<i>P. aeruginosa</i>	pchF	pyochelin biosynthesis	176/189	58/305	PA4226	pchE	dihydroaeruginosic acid synthetase	39.57	<i>Pseudomonas</i> sp. R2-7-07 9503	C4J86_RS180	amino acid adenylation	26

										60	domain-containing protein	
<i>P. aeruginosa</i>	plcB	motility	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	acp1	motility	189/189	0/305	PA2966	acpP	acyl carrier protein	53.4	<i>Pseudomonas mesoacidophila</i>	B7P44_RS05820	acyl carrier protein	61.11
<i>P. aeruginosa</i>	mucE	mucoidy	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	srfA	mucoidy	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	tse1	toxin/antitoxin operon	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	tsi1	toxin/antitoxin operon	189/189	1/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	tse2	toxin/antitoxin operon	187/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	tsi2	toxin/antitoxin operon	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	tse3	toxin/antitoxin operon	189/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	tsi3	toxin/antitoxin operon	186/189	0/305	-	-	-	-	-	-	-	-
<i>P. aeruginosa</i>	esrC	multidrug efflux	189/189	0/305	PA4600	nfxB	transcriptional regulator NfxB	60.71	<i>Pseudomonas chlororaphis</i>	C4K30_RS15575	TetR/AcrR family transcriptional regulator	66.47

<i>P. aeruginosa</i>	mexC	multidrug efflux	186/189	240/305	PA0425	mexA	resistance-nodulation cell division (RND) multidrug efflux membrane fusion protein MexA precursor	44.76	<i>P. sp</i> URMO17WK12 I11	PSHI_RS20595	hemolysin secretion protein D	27.78
<i>P. aeruginosa</i>	mexD	multidrug efflux	188/189	264/305	PA2018	mexY	resistance nodulation cell division (RND) multidrug efflux transporter MexY	52	<i>Pseudomonas alcaligenes</i>	A0T30_RS19090	multidrug efflux RND transporter permease subunit	53.09
<i>P. aeruginosa</i>	lecB	biofilm formation	188/189	0/305	-	-	-	-	-	-	-	-
<i>P. chlororaphis</i>	C4K22 RS23595	holin family	43/43	0/451	-	-	-	-	-	-	-	-
<i>P. chlororaphis</i>	C4K22 RS23415	mitomycin biosynthesis	42/43	0/451	-	-	-	-	-	-	-	-

A previous comparative analysis using the genomes of *Pseudomonas* species has demonstrated that MucE orthologues are found only in *P. aeruginosa* strains [81] this was confirmed in this study. SrfA appears to act as a stress response facilitator and positively regulates the expression of the *algD* operon. Furthermore, the sequence of SrfA was found to be extremely well conserved only in *P. aeruginosa*, suggesting that it is of high evolutionary importance in this species [83], again in agreement with our analysis.

Another very intriguing finding is the detection of three *P. aeruginosa*-specific toxin-antitoxin operons (Tse1-Tsi1, Tse2-Tsi2, Tse3-Tsi3) whose toxins are inserted into the periplasm of antagonizing bacteria by the H1 type VI (H1-T6SS) secretion system [84,85]. The Tse1 toxin (a peptidoglycan degrading amidase) is found in all *P. aeruginosa* strains, and in none of the other *Pseudomonas* strains, whereas its antitoxin Tsi1 is found in all *P. aeruginosa* strains, and in only one (*P. alcaligenes* NEB 585 4015) of the other *Pseudomonas* strains. The Tse2 toxin (an NAD-dependent cytotoxin) is found in 187/189 *P. aeruginosa* strains, and in none of the other *Pseudomonas* strains, whereas its antitoxin Tsi2 is found in all *P. aeruginosa* strains, and in none of the other *Pseudomonas* strains. The Tse3 toxin (peptidoglycan-degrading muramidase) is found in all *P. aeruginosa* strains, and in none of the other *Pseudomonas* strains, whereas its antitoxin Tsi3 is found in 186/189 of *P. aeruginosa* strains, and in none of the other *Pseudomonas* strains. It is reasonable to hypothesize that these three operons may allow *P. aeruginosa* to antagonize and prevail over other bacteria for the same ecological niche. Therefore, it would be interesting to investigate whether aerosols containing any of the three antitoxins (Tsi1-3) could block *P. aeruginosa* infections in CF patients by protecting other antagonizing bacteria.

EsrC encodes a negative regulator (moderator) of the *mexC/D-oprJ* envelope stress-inducible multidrug efflux operon that contains five cistrons. This operon is suggested to play a role in the remodeling of *P. aeruginosa* membranes in response to envelope stress [86]. *EsrC* is present in all *P. aeruginosa* strains, whereas the whole operon is found intact in 183/189 strains. However, this regulator (*EsrC*) that moderates (but does not totally inhibit) the expression of this efflux pump system is not detected in any other strains of the genus, although *mexC* and *mexD* are widely present (in 87% and 79%, respectively). Therefore, we suggest that this particular response (moderate expression of the efflux system) that affects membrane remodeling is a fine-tuning adaptation, specific for *P. aeruginosa*, which requires *EsrC*.

We also identified 19 other orthologous protein groups that are present in at least 188/189 *P. aeruginosa* strains and absent in all the other strains of the genus. Sixteen of them were annotated as hypothetical; however, of particular interest is fucose-binding lectin PA-III (lecB: PA3361). This lectin is important for biofilm formation and is believed to facilitate bacterial adhesion to the airway mucosa [87–89].

3.4. *Pseudomonas chlororaphis*-Specific Core Proteins with an Important Role in Niche Adaptation

P. chlororaphis can be employed as a biocontrol agent (used as inoculant) in agriculture against certain fungal plant pathogens via production of phenazine-type antibiotics [90]. In this study, a phage-like protein (C4K22 RS23595), belonging to the holin family, has been detected as a *P. chlororaphis*-specific core protein. Members of the holin family are small transmembrane proteins that accumulate in the bacterial cell membrane and permeabilize it, thus allowing the endolysins produced by certain phages to reach and degrade the peptidoglycan layer [91]. Recently, it was demonstrated that the holin gene is present in a *P. chlororaphis* strain that produces distinct high molecular weight bacteriocins. These bacteriocins resemble bacteriophage tails known as R-type tailocins [92] and help *P. chlororaphis* to outcompete closely related bacteria and to persist in rhizosphere communities [92,93].

Furthermore, our analysis identified another interesting protein-coding gene (C4K22 RS23415). It was annotated as a mitomycin-biosynthesis protein and had orthologues in all but one *P. chlororaphis* strains and did not have any orthologues in the other *Pseudomonas* proteomes. Mitomycin is an anticancer drug produced by *Streptomyces* spp. Recently, it has been demonstrated that mitomycin in combination with a tobramycin-ciprofloxacin hybrid is very effective (at least in vitro) against multidrug-resistant Gram-negative bacteria [94]. Although mitomycin production by *P. chlororaphis* has not yet been demonstrated, it is plausible that the compound produced by C4K22 RS23415 is mitomycin, or something similar. Thus, we speculate that it may confer a distinct advantage over other Gram-negative root colonizers and may represent an adaptation by *P. chlororaphis* to its ecological niche.

4. Conclusions and Future Perspectives

The *Pseudomonas* genus is very diverse and includes organisms from many different habitats that may function as human, animal, or plant pathogens, but they may also promote plant growth and nitrogen fixation [1,2,95]. This analysis of more than 490 complete *Pseudomonas* proteomes revealed six major evolutionary groups that each have sufficient numbers of genomes/proteomes for core proteome analyses [17]. We identified the core protein sets of the entire genus and each of the six evolutionary groups, as well as the properties of each core set. However, we also observed many species with a very low number of sequenced strains. In addition, due to the high number of proteomes from *P. chlororaphis* and especially from *P. aeruginosa*, our computational approaches identified several proteins that may play a significant role in species-specific adaptations. Although the number of *P. aeruginosa* proteomes was sufficient to provide a saturated core set, this was not the case for the other five evolutionary groups. In addition, although *P. chlororaphis* is a well-defined species, the other four groups of *P. fluorescens*, *P. putida*, *P. stutzeri*, and *P. syringae* harbor much more genetic diversity, and each one of these evolutionary groups is not actually one species, rather, each group includes more than one species. In the near future, sufficient numbers of high-quality genomes from strains of the same species should become available for many species of the genus. Thus, this type of computational analysis clearly demonstrates that the secrets of species-specific adaptation events for many pseudomonads will be unraveled in the near future, as many of their strains will be sequenced.

Supplementary Materials: The following are available online at www.mdpi.com/1424-2818/12/8/289/s1, Figure S1: Phylogenomic trees of the various *Pseudomonas* groups and the genus, Table S1: The genus core proteins. Table S2: Gene ontology of the genus core4. Table S3: The *P. aeruginosa* core proteins. Table S4: The *P. chlororaphis* core proteins. Table S5: The *P. stutzeri* core proteins. Table S6: The *P. putida* core proteins. Table S7: The *P. fluorescens* core proteins. Table S8: The *P. syringae* core proteins. Table S9: Information about the various groups. Table S10: Core proteins that are also essential. Table S11: Comparison with Orthofinder.

Author Contributions: Conceptualization, G.D.A., S.G.O., D.M.; methodology, G.D.A.; software, M.N., G.D.A.; formal analysis, M.N., D.M., S.G.O., G.D.A.; writing—original draft preparation, M.N., D.M., S.G.O., G.D.A.; supervision, G.D.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: M. Nikolaidis would like to thank the Bodossakis foundation for financial support (MSC studentship: BDA-394).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Silby, M.W.; Winstanley, C.; Godfrey, S.A.C.; Levy, S.B.; Jackson, R.W. *Pseudomonas* genomes: Diverse and adaptable. *FEMS Microbiol. Rev.* **2011**, *35*, 652–680, doi:10.1111/j.1574-6976.2011.00269.x.
2. Peix, A.; Ramírez-Bahena, M.-H.; Velázquez, E. The current status on the taxonomy of *Pseudomonas* revisited: An update. *Infect. Genet. Evol.* **2018**, *57*, 106–116, doi:10.1016/j.meegid.2017.10.026.
3. Moradali, M.F.; Ghods, S.; Rehm, B.H.A. *Pseudomonas aeruginosa* Lifestyle: A Paradigm for Adaptation, Survival, and Persistence. *Front. Cell Infect. Microbiol.* **2017**, *7*, 39, doi:10.3389/fcimb.2017.00039.
4. Lalucat, J.; Bennasar, A.; Bosch, R.; García-Valdés, E.; Palleroni, N.J. Biology of *Pseudomonas stutzeri*. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 510–547, doi:10.1128/MMBR.00047-05.
5. Xin, X.-F.; Kvitko, B.; He, S.Y. *Pseudomonas syringae*: What it takes to be a pathogen. *Nat. Rev. Microbiol.* **2018**, *16*, 316–328, doi:10.1038/nrmicro.2018.17.
6. Mulet, M.; Lalucat, J.; García-Valdés, E. DNA sequence-based analysis of the *Pseudomonas* species. *Environ. Microbiol.* **2010**, *12*, 1513–1530, doi:10.1111/j.1462-2920.2010.02181.x.
7. Scales, B.S.; Dickson, R.P.; LiPuma, J.J.; Huffnagle, G.B. Microbiology, genomics, and clinical significance of the *Pseudomonas fluorescens* species complex, an unappreciated colonizer of humans. *Clin. Microbiol. Rev.* **2014**, *27*, 927–948, doi:10.1128/CMR.00044-14.
8. Anderson, J.A.; Staley, J.; Challender, M.; Heuton, J. Safety of *Pseudomonas chlororaphis* as a gene source for genetically modified crops. *Transgenic Res.* **2018**, *27*, 103–113, doi:10.1007/s11248-018-0061-6.
9. Iyer, R.; Damania, A. Draft Genome Sequence of *Pseudomonas putida* CBF10-2, a Soil Isolate with Bioremediation Potential in Agricultural and Industrial Environmental Settings. *Genome Announc.* **2016**, *4*, doi:10.1128/genomeA.00670-16.
10. Papadopoulou, E.S.; Perruchon, C.; Vasileiadis, S.; Rousidou, C.; Tanou, G.; Samiotaki, M.; Molassiotis, A.; Karpouzas, D.G. Metabolic and Evolutionary Insights in the Transformation of Diphenylamine by a *Pseudomonas putida* Strain Unravelling by Genomic, Proteomic, and Transcription Analysis. *Front. Microbiol.* **2018**, *9*, 676, doi:10.3389/fmicb.2018.00676.
11. Anzai, Y.; Kim, H.; Park, J.Y.; Wakabayashi, H.; Oyaizu, H. Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence. *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 1563–1589, doi:10.1099/00207713-50-4-1563.
12. Woese, C.R.; Stackebrandt, E.; Weisburg, W.G.; Paster, B.J.; Madigan, M.T.; Fowler, V.J.; Hahn, C.M.; Blanz, P.; Gupta, R.; Nealson, K.H.; et al. The phylogeny of purple bacteria: The alpha subdivision. *Syst. Appl. Microbiol.* **1984**, *5*, 315–326, doi:10.1016/s0723-2020(84)80034-x.
13. Özen, A.I.; Ussery, D.W. Defining the *Pseudomonas* genus: Where do we draw the line with *Azotobacter*? *Microb. Ecol.* **2012**, *63*, 239–248, doi:10.1007/s00248-011-9914-8.
14. Tindall, B.J.; Rosselló-Móra, R.; Busse, H.-J.; Ludwig, W.; Kämpfer, P. Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol.* **2010**, *60*, 249–266, doi:10.1099/ijs.0.016949-0.
15. Palleroni, N.J. *Pseudomonas*. In *Bergey's Manual of Systematics of Archaea and Bacteria*; Whitman, W.B., Rainey, F., Kämpfer, P., Trujillo, M., Chun, J., DeVos, P., Hedlund, B., Dedysh, S., Eds.; John Wiley & Sons, Ltd.: Chichester, UK, 2015; p. 1, ISBN 978-1-118-96060-8.
16. Bennasar, A.; Mulet, M.; Lalucat, J.; García-Valdés, E. PseudoMLSA: A database for multigenic sequence analysis of *Pseudomonas* species. *BMC Microbiol.* **2010**, *10*, 118, doi:10.1186/1471-2180-10-118.
17. Vernikos, G.; Medini, D.; Riley, D.R.; Tettelin, H. Ten years of pan-genome analyses. *Curr. Opin. Microbiol.* **2015**, *23*, 148–154, doi:10.1016/j.mib.2014.11.016.
18. García-Valdés, E.; Lalucat, J. *Pseudomonas*: Molecular Phylogeny and Current Taxonomy. In *Pseudomonas: Molecular and Applied Biology*; Kahlon, R.S., Ed.; Springer International Publishing: Cham, Switzerland, 2016; pp. 1–23, ISBN 978-3-319-31197-5.
19. Lalucat, J.; Mulet, M.; Gomila, M.; García-Valdés, E. Genomics in Bacterial Taxonomy: Impact on the Genus *Pseudomonas*. *Genes* **2020**, *11*, 139, doi:10.3390/genes11020139.

20. Jolley, K.A.; Bliss, C.M.; Bennett, J.S.; Bratcher, H.B.; Brehony, C.; Colles, F.M.; Wimalarathna, H.; Harrison, O.B.; Sheppard, S.K.; Cody, A.J.; et al. Ribosomal multilocus sequence typing: Universal characterization of bacteria from domain to strain. *Microbiology* **2012**, *158*, 1005–1015, doi:10.1099/mic.0.055459-0.
21. Kyrpides, N.C.; Hugenholtz, P.; Eisen, J.A.; Woyke, T.; Göker, M.; Parker, C.T.; Amann, R.; Beck, B.J.; Chain, P.S.G.; Chun, J.; et al. Genomic encyclopedia of bacteria and archaea: Sequencing a myriad of type strains. *PLoS Biol.* **2014**, *12*, e1001920, doi:10.1371/journal.pbio.1001920.
22. Whitman, W.B.; Woyke, T.; Klenk, H.-P.; Zhou, Y.; Lilburn, T.G.; Beck, B.J.; De Vos, P.; Vandamme, P.; Eisen, J.A.; Garrity, G.; et al. Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase III: The genomes of soil and plant-associated and newly described type strains. *Stand. Genom. Sci.* **2015**, *10*, 26, doi:10.1186/s40793-015-0017-x.
23. Wu, D.; Hugenholtz, P.; Mavromatis, K.; Pukall, R.; Dalin, E.; Ivanova, N.N.; Kunin, V.; Goodwin, L.; Wu, M.; Tindall, B.J.; et al. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* **2009**, *462*, 1056–1060, doi:10.1038/nature08656.
24. Gogarten, J.P.; Townsend, J.P. Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* **2005**, *3*, 679–687, doi:10.1038/nrmicro1204.
25. Kunin, V.; Ouzounis, C.A. The balance of driving forces during genome evolution in prokaryotes. *Genome Res.* **2003**, *13*, 1589–1594, doi:10.1101/gr.1092603.
26. Kunin, V.; Goldovsky, L.; Darzentas, N.; Ouzounis, C.A. The net of life: Reconstructing the microbial phylogenetic network. *Genome Res.* **2005**, *15*, 954–959, doi:10.1101/gr.3666505.
27. Parks, D.H.; Chuvochina, M.; Waite, D.W.; Rinke, C.; Skarshewski, A.; Chaumeil, P.-A.; Hugenholtz, P. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **2018**, *36*, 996–1004, doi:10.1038/nbt.4229.
28. Tettelin, H.; Massignani, V.; Cieslewicz, M.J.; Donati, C.; Medini, D.; Ward, N.L.; Angiuoli, S.V.; Crabtree, J.; Jones, A.L.; Durkin, A.S.; et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome”. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13950–13955, doi:10.1073/pnas.0506758102.
29. Jun, S.-R.; Wassenaar, T.M.; Nookaew, I.; Hauser, L.; Wanchai, V.; Land, M.; Timm, C.M.; Lu, T.-Y.S.; Schadt, C.W.; Doktycz, M.J.; et al. Diversity of *Pseudomonas* Genomes, Including *Populus*-Associated Isolates, as Revealed by Comparative Genome Analysis. *Appl. Environ. Microbiol.* **2016**, *82*, 375–383, doi:10.1128/AEM.02612-15.
30. Winsor, G.L.; Griffiths, E.J.; Lo, R.; Dhillon, B.K.; Shay, J.A.; Brinkman, F.S.L. Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* **2016**, *44*, D646–D653, doi:10.1093/nar/gkv1227.
31. Freschi, L.; Jeukens, J.; Kukavica-Ibrulj, I.; Boyle, B.; Dupont, M.-J.; Laroche, J.; Larose, S.; Maaroufi, H.; Fothergill, J.L.; Moore, M.; et al. Clinical utilization of genomics data produced by the international *Pseudomonas aeruginosa* consortium. *Front. Microbiol.* **2015**, *6*, 1036, doi:10.3389/fmicb.2015.01036.
32. Jeukens, J.; Emond-Rheault, J.-G.; Freschi, L.; Kukavica-Ibrulj, I.; Levesque, R.C. Major Release of 161 Whole-Genome Sequences from the International *Pseudomonas* Consortium Database. *Microbiol. Resour. Announc.* **2019**, *8*, doi:10.1128/MRA.00013-19.
33. Konstantinidis, K.T.; Tiedje, J.M. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2567–2572, doi:10.1073/pnas.0409727102.
34. Richter, M.; Rosselló-Móra, R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 19126–19131, doi:10.1073/pnas.0906412106.
35. Tran, P.N.; Savka, M.A.; Gan, H.M. In-silico Taxonomic Classification of 373 Genomes Reveals Species Misidentification and New Genospecies within the Genus *Pseudomonas*. *Front. Microbiol.* **2017**, *8*, 1296, doi:10.3389/fmicb.2017.01296.

36. Bentley, S.D.; Vernikos, G.S.; Snyder, L.A.S.; Churcher, C.; Arrowsmith, C.; Chillingworth, T.; Cronin, A.; Davis, P.H.; Holroyd, N.E.; Jagels, K.; et al. Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS Genet.* **2007**, *3*, e23, doi:10.1371/journal.pgen.0030023.
37. Hiller, N.L.; Janto, B.; Hogg, J.S.; Boissy, R.; Yu, S.; Powell, E.; Keefe, R.; Ehrlich, N.E.; Shen, K.; Hayes, J.; et al. Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: Insights into the pneumococcal supragenome. *J. Bacteriol.* **2007**, *189*, 8186–8195, doi:10.1128/JB.00690-07.
38. Méric, G.; Yahara, K.; Mageiros, L.; Pascoe, B.; Maiden, M.C.J.; Jolley, K.A.; Sheppard, S.K. A reference pan-genome approach to comparative bacterial genomics: Identification of novel epidemiological markers in pathogenic *Campylobacter*. *PLoS ONE* **2014**, *9*, e92798, doi:10.1371/journal.pone.0092798.
39. Rasko, D.A.; Rosovitz, M.J.; Myers, G.S.A.; Mongodin, E.F.; Fricke, W.F.; Gajer, P.; Crabtree, J.; Sebaihia, M.; Thomson, N.R.; Chaudhuri, R.; et al. The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* **2008**, *190*, 6881–6893, doi:10.1128/JB.00619-08.
40. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797, doi:10.1093/nar/gkh340.
41. Gouy, M.; Guindon, S.; Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **2010**, *27*, 221–224, doi:10.1093/molbev/msp259.
42. Pritchard, L.; Glover, R.H.; Humphris, S.; Elphinstone, J.G.; Toth, I.K. Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods* **2016**, *8*, 12–24, doi:10.1039/C5AY02550H.
43. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238, doi:10.1186/s13059-019-1832-y.
44. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60, doi:10.1038/nmeth.3176.
45. Huerta-Cepas, J.; Szklarczyk, D.; Forslund, K.; Cook, H.; Heller, D.; Walter, M.C.; Rattei, T.; Mende, D.R.; Sunagawa, S.; Kuhn, M.; et al. eggNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **2016**, *44*, D286–293, doi:10.1093/nar/gkv1248.
46. Moriya, Y.; Itoh, M.; Okuda, S.; Yoshizawa, A.C.; Kanehisa, M. KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **2007**, *35*, W182–W185, doi:10.1093/nar/gkm321.
47. Galperin, M.Y.; Makarova, K.S.; Wolf, Y.I.; Koonin, E.V. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* **2015**, *43*, D261–D269, doi:10.1093/nar/gku1223.
48. Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; et al. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **2000**, *406*, 959–964, doi:10.1038/35023079.
49. Loveridge, E.J.; Jones, C.; Bull, M.J.; Moody, S.C.; Kahl, M.W.; Khan, Z.; Neilson, L.; Tomeva, M.; Adams, S.E.; Wood, A.C.; et al. Reclassification of the Specialized Metabolite Producer *Pseudomonas mesoacidophila* ATCC 31433 as a Member of the *Burkholderia cepacia* Complex. *J. Bacteriol.* **2017**, *199*, doi:10.1128/JB.00125-17.
50. Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X.; Thomas, P.D. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **2019**, *47*, D419–D426, doi:10.1093/nar/gky1038.
51. Hesse, C.; Schulz, F.; Bull, C.T.; Shaffer, B.T.; Yan, Q.; Shapiro, N.; Hassan, K.A.; Varghese, N.; Elbourne, L.D.H.; Paulsen, I.T.; et al. Genome-based evolutionary history of *Pseudomonas* spp. *Environ. Microbiol.* **2018**, *20*, 2142–2159, doi:10.1111/1462-2920.14130.

52. Gomila, M.; Peña, A.; Mulet, M.; Lalucat, J.; García-Valdés, E. Phylogenomics and systematics in *Pseudomonas*. *Front. Microbiol.* **2015**, *6*, 214, doi:10.3389/fmicb.2015.00214.
53. Ozer, E.A.; Allen, J.P.; Hauser, A.R. Characterization of the core and accessory genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGent. *BMC Genomics* **2014**, *15*, 737, doi:10.1186/1471-2164-15-737.
54. Subedi, D.; Vijay, A.K.; Kohli, G.S.; Rice, S.A.; Willcox, M. Comparative genomics of clinical strains of *Pseudomonas aeruginosa* strains isolated from different geographic sites. *Sci. Rep.* **2018**, *8*, 15668, doi:10.1038/s41598-018-34020-7.
55. Freschi, L.; Vincent, A.T.; Jeukens, J.; Emond-Rheault, J.-G.; Kukavica-Ibrulj, I.; Dupont, M.-J.; Charette, S.J.; Boyle, B.; Levesque, R.C. The *Pseudomonas aeruginosa* Pan-Genome Provides New Insights on Its Population Structure, Horizontal Gene Transfer, and Pathogenicity. *Genome Biol. Evol.* **2019**, *11*, 109–120, doi:10.1093/gbe/evy259.
56. Valot, B.; Guyeux, C.; Rolland, J.Y.; Mazouzi, K.; Bertrand, X.; Hocquet, D. What It Takes to Be a *Pseudomonas aeruginosa*? The Core Genome of the Opportunistic Pathogen Updated. *PLoS ONE* **2015**, *10*, e0126468, doi:10.1371/journal.pone.0126468.
57. Weiser, R.; Green, A.E.; Bull, M.J.; Cunningham-Oakes, E.; Jolley, K.A.; Maiden, M.C.J.; Hall, A.J.; Winstanley, C.; Weightman, A.J.; Donoghue, D.; et al. Not all *Pseudomonas aeruginosa* are equal: Strains from industrial sources possess uniquely large multireplicon genomes. *Microb. Genom.* **2019**, *5*, doi:10.1099/mgen.0.000276.
58. Poulsen, B.E.; Yang, R.; Clatworthy, A.E.; White, T.; Osmulski, S.J.; Li, L.; Penaranda, C.; Lander, E.S.; Shores, N.; Hung, D.T. Defining the core essential genome of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 10072–10080, doi:10.1073/pnas.1900570116.
59. Liberati, N.T.; Urbach, J.M.; Miyata, S.; Lee, D.G.; Drenkard, E.; Wu, G.; Villanueva, J.; Wei, T.; Ausubel, F.M. An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2833–2838, doi:10.1073/pnas.0511100103.
60. Skurnik, D.; Roux, D.; Aschard, H.; Cattoir, V.; Yoder-Himes, D.; Lory, S.; Pier, G.B. A comprehensive analysis of In Vitro and in vivo genetic fitness of *Pseudomonas aeruginosa* using high-throughput sequencing of transposon libraries. *PLoS Pathog.* **2013**, *9*, e1003582, doi:10.1371/journal.ppat.1003582.
61. Turner, K.H.; Wessel, A.K.; Palmer, G.C.; Murray, J.L.; Whiteley, M. Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 4110–4115, doi:10.1073/pnas.1419677112.
62. Lee, S.A.; Gallagher, L.A.; Thongdee, M.; Staudinger, B.J.; Lippman, S.; Singh, P.K.; Manoil, C. General and condition-specific essential functions of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5189–5194, doi:10.1073/pnas.1422186112.
63. Silby, M.W.; Cerdeño-Tárraga, A.M.; Vernikos, G.S.; Giddens, S.R.; Jackson, R.W.; Preston, G.M.; Zhang, X.-X.; Moon, C.D.; Gehrig, S.M.; Godfrey, S.A.C.; et al. Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biol.* **2009**, *10*, R51, doi:10.1186/gb-2009-10-5-r51.
64. Hernández-Salmerón, J.E.; Moreno-Hagelsieb, G.; Santoyo, G. Genome Comparison of *Pseudomonas fluorescens* UM270 with Related Fluorescent Strains Unveils Genes Involved in Rhizosphere Competence and Colonization. *J. Genomics* **2017**, *5*, 91–98, doi:10.7150/jgen.21588.
65. Garrido-Sanz, D.; Arrebola, E.; Martínez-Granero, F.; García-Méndez, S.; Muriel, C.; Blanco-Romero, E.; Martín, M.; Rivilla, R.; Redondo-Nieto, M. Classification of Isolates from the *Pseudomonas fluorescens* Complex into Phylogenomic Groups Based in Group-Specific Markers. *Front. Microbiol.* **2017**, *8*, 413, doi:10.3389/fmicb.2017.00413.
66. Udaondo, Z.; Molina, L.; Segura, A.; Duque, E.; Ramos, J.L. Analysis of the core genome and pangenome of *Pseudomonas putida*. *Environ. Microbiol.* **2016**, *18*, 3268–3283, doi:10.1111/1462-2920.13015.

67. Koehorst, J.J.; van Dam, J.C.J.; van Heck, R.G.A.; Saccenti, E.; Dos Santos, V.A.P.M.; Suarez-Diez, M.; Schaap, P.J. Comparison of 432 *Pseudomonas* strains through integration of genomic, functional, metabolic and expression data. *Sci. Rep.* **2016**, *6*, 38699, doi:10.1038/srep38699.
68. Lhospice, S.; Gomez, N.O.; Ouerdane, L.; Brutesco, C.; Ghssein, G.; Hajjar, C.; Liratni, A.; Wang, S.; Richaud, P.; Bleves, S.; et al. *Pseudomonas aeruginosa* zinc uptake in chelating environment is primarily mediated by the metallophore pseudopaline. *Sci. Rep.* **2017**, *7*, 17132, doi:10.1038/s41598-017-16765-9.
69. McFarlane, J.S.; Lamb, A.L. Biosynthesis of an Opine Metallophore by *Pseudomonas aeruginosa*. *Biochemistry* **2017**, *56*, 5967–5971, doi:10.1021/acs.biochem.7b00804.
70. Ghssein, G.; Brutesco, C.; Ouerdane, L.; Fojcik, C.; Izaute, A.; Wang, S.; Hajjar, C.; Lobinski, R.; Lemaire, D.; Richaud, P.; et al. Biosynthesis of a broad-spectrum nicotianamine-like metallophore in *Staphylococcus aureus*. *Science* **2016**, *352*, 1105–1109, doi:10.1126/science.aaf1018.
71. Crespo, A.; Pedraz, L.; Astola, J.; Torrents, E. *Pseudomonas aeruginosa* Exhibits Deficient Biofilm Formation in the Absence of Class II and III Ribonucleotide Reductases Due to Hindered Anaerobic Growth. *Front. Microbiol.* **2016**, *7*, 688, doi:10.3389/fmicb.2016.00688.
72. Mossialos, D.; Amoutzias, G.D. Role of siderophores in cystic fibrosis pathogenesis: Foes or friends? *Int. J. Med. Microbiol.* **2009**, *299*, 87–98, doi:10.1016/j.ijmm.2008.06.008.
73. Mossialos, D.; Amoutzias, G.D. Siderophores in fluorescent pseudomonads: New tricks from an old dog. *Future Microbiol.* **2007**, *2*, 387–395, doi:10.2217/17460913.2.4.387.
74. Amoutzias, G.D.; Van de Peer, Y.; Mossialos, D. Evolution and taxonomic distribution of nonribosomal peptide and polyketide synthases. *Future Microbiol.* **2008**, *3*, 361–370, doi:10.2217/17460913.3.3.361.
75. Amoutzias, G.D.; Chaliotis, A.; Mossialos, D. Discovery Strategies of Bioactive Compounds Synthesized by Nonribosomal Peptide Synthetases and Type-I Polyketide Synthases Derived from Marine Microbiomes. *Mar. Drugs* **2016**, *14*, doi:10.3390/md14040080.
76. Mossialos, D.; Ochsner, U.; Baysse, C.; Chablain, P.; Pirnay, J.-P.; Koedam, N.; Budzikiewicz, H.; Fernández, D.U.; Schäfer, M.; Ravel, J.; et al. Identification of new, conserved, non-ribosomal peptide synthetases from fluorescent pseudomonads involved in the biosynthesis of the siderophore pyoverdine. *Mol. Microbiol.* **2002**, *45*, 1673–1685, doi:10.1046/j.1365-2958.2002.03120.x.
77. Barker, A.P.; Vasil, A.I.; Filloux, A.; Ball, G.; Wilderman, P.J.; Vasil, M.L. A novel extracellular phospholipase C of *Pseudomonas aeruginosa* is required for phospholipid chemotaxis. *Mol. Microbiol.* **2004**, *53*, 1089–1098, doi:10.1111/j.1365-2958.2004.04189.x.
78. Wagner, V.E.; Bushnell, D.; Passador, L.; Brooks, A.I.; Iglewski, B.H. Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: Effects of growth phase and environment. *J. Bacteriol.* **2003**, *185*, 2080–2095, doi:10.1128/jb.185.7.2080-2095.2003.
79. Meyer, K.C.; Sharma, A.; Brown, R.; Weatherly, M.; Moya, F.R.; Lewandoski, J.; Zimmerman, J.J. Function and composition of pulmonary surfactant and surfactant-derived fatty acid profiles are altered in young adults with cystic fibrosis. *Chest* **2000**, *118*, 164–174, doi:10.1378/chest.118.1.164.
80. Ma, J.-C.; Wu, Y.-Q.; Cao, D.; Zhang, W.-B.; Wang, H.-H. Only Acyl Carrier Protein 1 (AcpP1) Functions in *Pseudomonas aeruginosa* Fatty Acid Synthesis. *Front. Microbiol.* **2017**, *8*, 2186, doi:10.3389/fmicb.2017.02186.
81. Qiu, D.; Eisinger, V.M.; Rowen, D.W.; Yu, H.D. Regulated proteolysis controls mucoid conversion in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8107–8112, doi:10.1073/pnas.0702660104.
82. Yin, Y.; Damron, F.H.; Withers, T.R.; Pritchett, C.L.; Wang, X.; Schurr, M.J.; Yu, H.D. Expression of mucoid induction factor MucE is dependent upon the alternate sigma factor AlgU in *Pseudomonas aeruginosa*. *BMC Microbiol.* **2013**, *13*, 232, doi:10.1186/1471-2180-13-232.
83. Wood, L.F.; Ohman, D.E. Cell wall stress activates expression of a novel stress response facilitator (SrfA) under σ^{22} (AlgT/U) control in *Pseudomonas aeruginosa*. *Microbiology* **2015**, *161*, 30–40, doi:10.1099/mic.0.081182-0.

84. Hood, R.D.; Singh, P.; Hsu, F.; Güvener, T.; Carl, M.A.; Trinidad, R.R.S.; Silverman, J.M.; Ohlson, B.B.; Hicks, K.G.; Plemel, R.L.; et al. A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe* **2010**, *7*, 25–37, doi:10.1016/j.chom.2009.12.007.
85. Russell, A.B.; Hood, R.D.; Bui, N.K.; LeRoux, M.; Vollmer, W.; Mougous, J.D. Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **2011**, *475*, 343–347, doi:10.1038/nature10244.
86. Purcell, A.; Fruci, M.; Mikalauskas, A.; Gilmour, C.; Poole, K. EsrC, an envelope stress-regulated repressor of the mexCD-oprJ multidrug efflux operon in *Pseudomonas aeruginosa*. *Environ. Microbiol.* **2015**, *17*, 186–198, doi:10.1111/1462-2920.12602.
87. Tielker, D.; Hacker, S.; Loris, R.; Strathmann, M.; Wingender, J.; Wilhelm, S.; Rosenau, F.; Jaeger, K.-E. *Pseudomonas aeruginosa* lectin LecB is located in the outer membrane and is involved in biofilm formation. *Microbiology* **2005**, *151*, 1313–1323, doi:10.1099/mic.0.27701-0.
88. Mitchell, E.; Houles, C.; Sudakevitz, D.; Wimmerova, M.; Gautier, C.; Pérez, S.; Wu, A.M.; Gilboa-Garber, N.; Imberty, A. Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients. *Nat. Struct. Mol. Biol.* **2002**, *9*, 918–921, doi:10.1038/nsb865.
89. Gilboa-Garber, N.; Katcoff, D.J.; Garber, N.C. Identification and characterization of *Pseudomonas aeruginosa* PA-III lectin gene and protein compared to PA-IL. *FEMS Immunol. Med. Microbiol.* **2000**, *29*, 53–57, doi:10.1111/j.1574-695X.2000.tb01505.x.
90. Chin-A-Woeng, T.F.; Bloembergen, G.V.; Mulders, I.H.; Dekkers, L.C.; Lugtenberg, B.J. Root colonization by phenazine-1-carboxamide-producing bacterium *Pseudomonas chlororaphis* PCL1391 is essential for biocontrol of tomato foot and root rot. *Mol. Plant Microbe Interact.* **2000**, *13*, 1340–1345, doi:10.1094/MPMI.2000.13.12.1340.
91. Wang, I.N.; Smith, D.L.; Young, R. Holins: The protein clocks of bacteriophage infections. *Annu. Rev. Microbiol.* **2000**, *54*, 799–825, doi:10.1146/annurev.micro.54.1.799.
92. Dorosky, R.J.; Yu, J.M.; Pierson, L.S.; Pierson, E.A. *Pseudomonas chlororaphis* Produces Two Distinct R-Tailocins That Contribute to Bacterial Competition in Biofilms and on Roots. *Appl. Environ. Microbiol.* **2017**, *83*, doi:10.1128/AEM.00706-17.
93. Dorosky, R.J.; Pierson, L.S.; Pierson, E.A. *Pseudomonas chlororaphis* Produces Multiple R-Tailocin Particles That Broaden the Killing Spectrum and Contribute to Persistence in Rhizosphere Communities. *Appl. Environ. Microbiol.* **2018**, *84*, doi:10.1128/AEM.01230-18.
94. Domalaon, R.; Ammeter, D.; Brizuela, M.; Gorityala, B.K.; Zhanel, G.G.; Schweizer, F. Repurposed Antimicrobial Combination Therapy: Tobramycin-Ciprofloxacin Hybrid Augments Activity of the Anticancer Drug Mitomycin C Against Multidrug-Resistant Gram-Negative Bacteria. *Front. Microbiol.* **2019**, *10*, 1556, doi:10.3389/fmicb.2019.01556.
95. Li, H.-B.; Singh, R.K.; Singh, P.; Song, Q.-Q.; Xing, Y.-X.; Yang, L.-T.; Li, Y.-R. Genetic Diversity of Nitrogen-Fixing and Plant Growth Promoting *Pseudomonas* Species Isolated from Sugarcane Rhizosphere. *Front. Microbiol.* **2017**, *8*, doi:10.3389/fmicb.2017.01268.

